Materials and Methods

RNA sequencing

ALI-NHBE cells were treated with 100 ng/ml flagellin for 3 hours and control cells were used for transcriptome analysis. Total RNAs were fragmented and converted into a cDNA library constructed with the protocol for paired-end sequencing with Illumina TruSeq RNA Sample Preparation Kit v2. The sequencing procedure of RNA-seq was conducted by Macrogen Inc (Seoul, Korea). The protocol consisted of RNA isolation, purification, RNA fragmentation, reverse transcription into cDNA, adding sequencing adapters, PCR amplification, and sequencing. The following steps were pre-processing and data analysis. Quality control of raw reads was performed, producing overall read quality, total bases, total reads, and GC (%). Adaptor sequences, contaminant DNA, and PCR duplicates were excluded to reduce bias in the analysis results. The RNA-seq reads were mapped to the reference genome using the TopHat and determined using Cufflinks which assembles transcripts, estimates transcript length and depth of coverage, and tests for differential expression. The transcript expression was quantified in FPKM (Fragments Per Kilobase of transcript per Million mapped reads) and estimated FC. Expression of the transcript that satisfied with a $|FC \geq 2|$ was considered as DEGs.

In order to explore the function and the pathways of the significant genes, the Gene - Enrichment and Functional Annotation Analysis (Gene Ontology using g:Profiler, http://biit.cs.ut.ee/gprofiler/) and KEGG Pathway Analysis (KEGG, http://www.genome.jp/kegg/pathway.html) were performed.

Statistical analysis

Except for FPKM (raw signal), values were log normalized, and FC was the converted value of the calculated log2 FC value to a linear scale, calculated as a power of 2. Volume was defined as a geometric mean of the expression values of the two groups and is calculated as square root

(control normalized value x test normalized value). The p-value for differential expression was extracted using Fisher's exact test, and corrected accounting false discovery rate (FDR).

Results

RNA sequencing and analysis of differentially expressed genes (DEGs)

To investigate differentiated NHBE cells' response to TLR5 ligand flagellin, transcriptional changes to flagellin were assessed in treated and control cells. The raw sequencing data has been deposited in NCBI sequence read archive (SRA) under the accession number PRJNA792916. In RNA-seq, a total of 76,339,980 of raw reads and 75,965,372 of clean reads were generated in control sample, and 83,725,824 of raw reads and 83,358,706 of clean reads were produced in treated sample. For each sample, the G20 and Q30 were 97% and 95%, respectively. The analysis revealed upregulated 60 genes and downregulated 28 genes (Fig. S1A and Table S2). Among the upregulated flagellin-responsive genes, the genes encoding chemokines (*CXCL5*, *CCL5*, *CXCL10*, *CXCL11*), matrix metalloproteinases (*MMP13*, *MMP9*, *MMP7*), and antimicrobial biomolecules (*DEFB4A*, *DEFB4B*, *MUC4*) were frequently observed. The difference in gene expression level expressed by log2 fold change (FC) against the volume of the expression value is presented in Fig. S1B. The top 5 ranking genes were *TINAGL1*, *SERPINA3*, *IGFBP3*, *KRT17*, and *KRT13*.

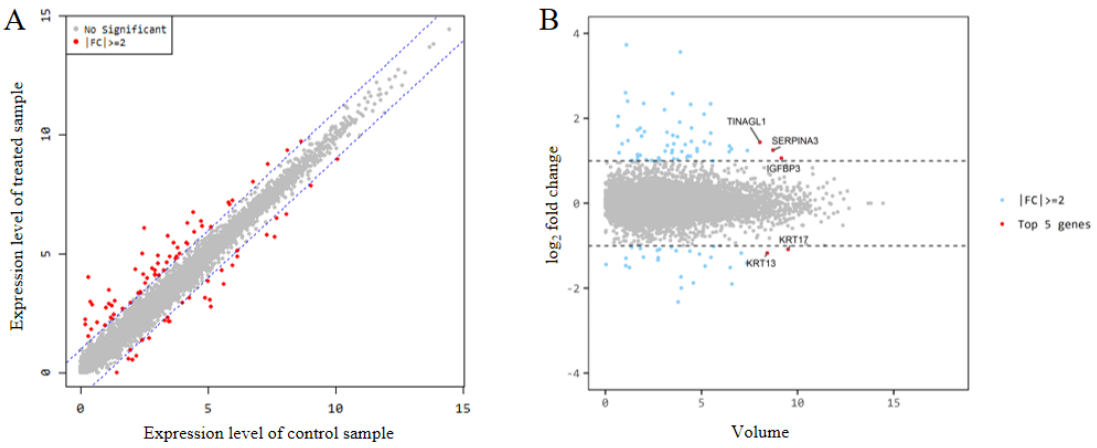

Fig. S1. Plots of RNA-sequencing gene profile altered by flagellin stimulation in primary human bronchial epithelial cells. A. Scatter plot shows the correlation of the gene expression profiles between the control and stimulated cells. The statistically significant DEGs are represented by red, and genes that were not differentially expressed are in grey. B. Volume plot shows the differences in gene expression level between control and stimulated groups. The X-axis shows the mean of the expression values of the two groups, and Y-axis displays the log2 fold changes values. Up- and downregulated transcripts are represented by blue, and the top 5 ranking genes are shown by red.

Gene ontology (GO) enrichment analysis of DEGs

GO was applied to identify characteristic biological attributes of RNA-seq data. Separate GO enrichment analysis for up- and downregulated genes was performed. GO analysis results revealed that DEGs belonged to certain molecular functions, cellular components, and biological processes (Fig. S2). Upregulated DEGs were involved in total of 246 GO terms (adjusted p-value < 0.05) including cytokine-mediated signaling pathway (GO:0019221, 23 DEGs), response to cytokine (GO:0034097, 24 DEGs), and cellular response to cytokine stimulus (GO:0071345, 23 DEGs) in biological process, extracellular space (GO:0005615, 30 DEGs), extracellular region (GO:0005576, 32 DEGs), and external encapsulating structure

(GO:0030312, 9 DEGs) in cellular components, and cytokine activity (GO:0005125, 8 DEGs), chemokine activity (GO:0008009, 4 DEGs), and CXCR chemokine receptor binding (GO:0045236, 3 DEGs) in molecular function. Downregulated DEGs were involved in 67 GO terms including cornification (GO:0070268, 4 DEGs) and animal organ development (GO:0048513, 14 DEGs) in biological process, intermediate filament (GO:0005882, 4 DEGs) and basal part of cell (GO:0045178, 4 DEGs) in cellular component, and pyridoxal phosphate binding (GO:0030170, 2 DEGs) and structural molecular activity (GO:0005198, 5 DEGs) in molecular function.

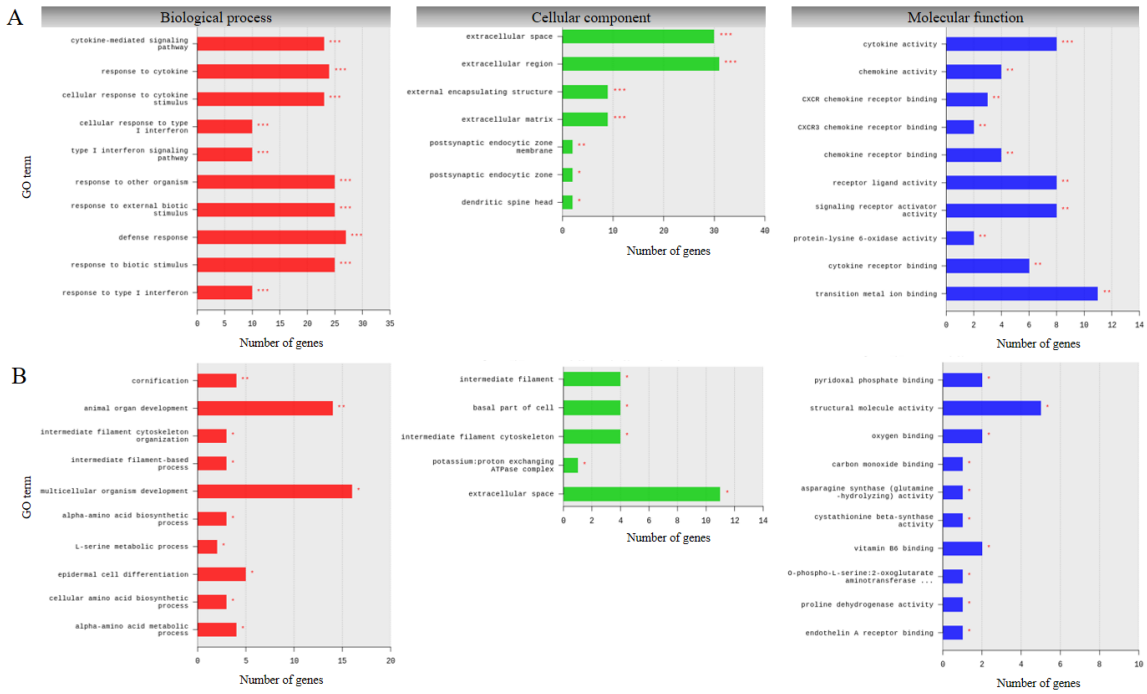

Fig. S2. GO enrichment analysis of the DEGs in flagellin stimulated versus control cells. A. Upregulated DEGs in flagellin-stimulated and control cells. B. Downregulated DEGs in stimulated and control cells. X-axis represents the number of DEG; Y-axis represents the GO terms names: red pillars represent biological process, green represents cellular component, and blue represents molecular function. Adjusted p-value < 0.05*, 0.01**, 0.001***.

Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis of DEGs

To estimate the number of DEGs contained at different class of KEGG pathways, pathway enrichment analysis was performed. KEGG pathway analysis showed that upregulated DEGs were mainly enriched in pathways such as IL-17 signaling pathway (8 DEGs), cytokine-cytokine receptor interaction (9 DEGs), and influenza A (7 DEGs), while downregulated DEGs were enriched in pathways such as *Staphylococcus aureus* infection (5 DEGs), estrogen signaling pathway (4 DEGs), and biosynthesis of amino acids (3 DEGs) (Fig. S3).

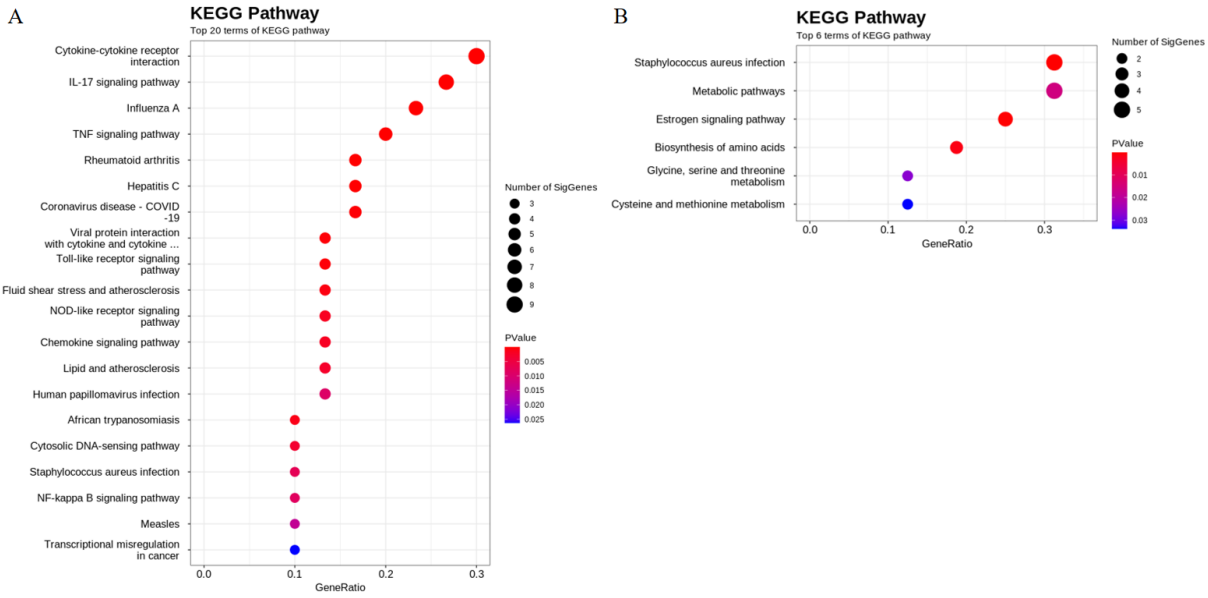

Fig. S3. KEGG enrichment analysis of the DEGs in flagellin stimulated versus control cells. A. KEGG pathways of upregulated DEGs in flagellin-stimulated and control NHBE cells. B. KEGG pathways of downregulated DEGs in stimulated and control NHBE cells. X-axis represents the gene ratio; Y-axis represents the KEGG pathway, the size of spots represents the number of genes, and the color of spots represents the p-value of enrichment. KEGG pathway analysis is adapted here from http://www.kegg.jp/kegg/kegg1.html.

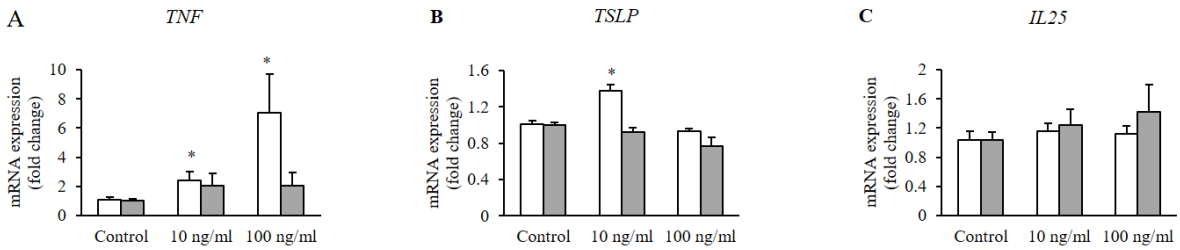

Fig. S4. mRNA expression of cytokines from flagellin stimulated NHBE cells. 3 hours, open bar; 24 hours, closed bar. The data are shown as the mean ± SEM from three independent experiments. *p value <0.05.

Table S1. List of primers sequences

Gene	Sequence
TNF	Forward 5'-CTTAGTGGGATACTCAGAACG-3'
	Reverse 5'-GGCGGTTCAGCCACTGGAGCT-3'
IL33	Forward 5'-GTGGAAGAACACAGCAAGCA-3'
	Reverse 5'-AAGGCAAAGCACTCCACAGT-3'
IL25	Forward 5'-CCAGGTGGTTGCATTCTTGG-3'
	Reverse 5'-TGGCTGTAGGTGTGGGTTCC-3'
TSLP	Forward 5'-CGCCTATGAGCAGCCACAT-3'
	Reverse 5'-CCGGCGGTGGGATTG-3'
CXCL8	Forward 5'-ACTGAGAGTGATTGAGAGTGGAC-3'
	Reverse 5'-AACCCTCTGCACCCAGTTTTC-3'
CSF2	Forward 5'-TTCTGCTTGTCATCCCCTTT-3'
	Reverse 5'-CTTCTGCCATGCCTGTATCA -3'
CXCL10	Forward 5'-GAAATTATTCCTGCAAGCCAATTT-3'
	Reverse 5'-TCACCCTTCTTTTTCATTGTAGCA-3'
CXCL11	Forward 5'-ATGAGTGTGAAGGGCATGGC-3'
	Reverse 5'-TCACTGCTTTTACCCCAGGG-3'
CXCL5	Forward 5'- AAGGTGGAAGTGGTAGCCTC-3'
	Reverse 5'-TCCTTGTTTCCACCGTCCAA-3'
CCL5	Forward 5'-CCTCGCTGTCATCCTCATTGCT-3'
	Reverse 5'-TACTCCCGAACCCATTTCTTCTC-3'
GAPDH	Forward 5'-TGGGCTACACTGAGCACCAG-3'
	Reverse 5'-GGGTGTCGCTGTTGAAGTCA-3'

Table S2. Full list of up- and downregulated genes in differentiated NHBE cells

Gene ID	Transcript ID	Gene	Gene description	Fold change
		Upregulated		
3627	NM_001565	*CXCL10*	chemokine (C-X-C motif) ligand 10	13.28
1592	NM_057157 NM_000783	*CYP26A1*	cytochrome P450, family 26, subfamily A, polypeptide 1	11.84
374897	NM_001166035 NM_001166034 NM_198538	*SBSN*	suprabasin	6.07
7018	NM_001063	*TF*	transferrin	6.00
91543	NM_080657	*RSAD2*	radical S-adenosyl methionine domain containing 2	5.30
6352	NM_002985 NM_001278736	*CCL5*	chemokine (C-C motif) ligand 5	5.09
84419	NM_197955 NM_032413	*C15orf48*	chromosome 15 open reading frame 48	5.06
3429	NM_001130080[a]	*IFI27*	interferon, alpha-inducible protein 27	5.02
5653	NM_001012964 NM_001012965 NM_002774	*KLK6*	kallikrein-related peptidase 6	4.61
2537	NM_002038 NM_022872 NM_022873	*IFI6*	interferon, alpha-inducible protein 6	4.28
6373	NM_001302123 NM_005409	*CXCL11*	chemokine (C-X-C motif) ligand 11	4.13
3620	NM_002164	*IDO1*	indoleamine 2,3-dioxygenase 1	3.76
200315	NM_001270406 NM_145699	*APOBEC3A*	apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3A	3.54
1673	NM_004942	*DEFB4A*	defensin, beta 4A	3.41
4600	NM_002463	*MX2*	myxovirus (influenza virus) resistance 2 (mouse)	3.40
11074	NM_007028	*TRIM31*	tripartite motif containing 31	3.30
3434	NM_001270930[b]	*IFIT1*	interferon-induced protein with tetratricopeptide repeats 1	3.22
100289462	NM_001205266	*DEFB4B*	defensin, beta 4B	3.19
9636	NM_005101	*ISG15*	ISG15 ubiquitin-like modifier	3.14
283422	NR_036555	*C12orf36*	chromosome 12 open reading frame 36	3.04

140686	NM_080614	*WFDC3*	WAP four-disulfide core domain 3	2.997004
4017	NM_002318	*LOXL2*	lysyl oxidase-like 2	2.964518
10141	NR_126517, NM_005750, NR_126519	*LINC01587*	long intergenic non-protein coding RNA 1587	2.789672
5650	NM_001207053, NM_001243126, NM_005046, NM_139277	*KLK7*	kallikrein-related peptidase 7	2.750379
4322	NM_002427	*MMP13*	matrix metallopeptidase 13 (collagenase 3)	2.731510
64129	NM_001204414, NM_001204415, NM_022164	*TINAGL1*	tubulointerstitial nephritis antigen-like 1	2.711534
84171	NM_032211	*LOXL4*	lysyl oxidase-like 4	2.643455
3290	NM_181755, NM_001206741, NM_005525	*HSD11B1*	hydroxysteroid (11-beta) dehydrogenase 1	2.643314
4050	NM_009588, NM_002341	*LTB*	lymphotoxin beta (TNF superfamily, member 3)	2.639311
64220	NM_001199042, NM_001199040, NM_022369, NM_001142619, NM_001142618, NM_001142617, NM_001142620, NM_001199041	*STRA6*	stimulated by retinoic acid 6	2.605982
55187	NM_018156, NM_015378	*VPS13D*	vacuolar protein sorting 13 homolog D (S. cerevisiae)	2.569915
10537	NM_006398	*UBD*	ubiquitin D	2.548432
3553	NM_000576	*IL1B*	interleukin 1, beta	2.437417
5327	NM_033011, NM_000930	*PLAT*	plasminogen activator, tissue	2.430821
12	NM_001085	*SERPINA3*	serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 3	2.387078
3635	NM_001017915, NM_005541	*INPP5D*	inositol polyphosphate-5-phosphatase, 145kDa	2.373419
7127	NM_006291	*TNFAIP2*	tumor necrosis factor, alpha-induced protein 2	2.365486
55601	NM_017631	*DDX60*	DEAD (Asp-Glu-Ala-Asp) box polypeptide 60	2.341334
3383	NM_000201	*ICAM1*	intercellular adhesion molecule 1	2.336481

4599	NM_002462, NM_001282920, NM_001178046, NM_001144925	*MX1*	myxovirus (influenza virus) resistance 1, interferon-inducible protein p78 (mouse)	2.333305
4318	NM_004994	*MMP9*	matrix metallopeptidase 9 (gelatinase B, 92kDa gelatinase, 92kDa type IV collagenase)	2.324859
54544	NM_019060	*CRCT1*	cysteine-rich C-terminal 1	2.285995
6374	NM_002994	*CXCL5*	chemokine (C-X-C motif) ligand 5	2.283614
121838	NR_026955	*LINC00284*	long intergenic non-protein coding RNA 284	2.216617
101927571	NR_110782	*LOC101927571*	n/a	2.192767
3437	NM_001289758, NM_001289759, NM_001549, NM_001031683	*IFIT3*	interferon-induced protein with tetratricopeptide repeats 3 isoform b	2.169077
8771	NM_003823	*TNFRSF6B*	tumor necrosis factor receptor superfamily, member 6b, decoy	2.157341
27189	NM_013278	*IL17C*	interleukin 17C	2.153270
4316	NM_002423	*MMP7*	matrix metallopeptidase 7 (matrilysin, uterine)	2.105077
26579	NM_001300924, NM_001300923, NM_001293291, NM_001293294, NM_001293296, NM_138768	*MYEOV*	myeloma-overexpressed gene protein isoform 2	2.102407
3486	NM_000598, NM_001013398	*IGFBP3*	insulin-like growth factor binding protein 3	2.087367
6623	NM_003087	*SNCG*	synuclein, gamma (breast cancer-specific protein 1)	2.067461
3433	NM_001547	*IFIT2*	interferon-induced protein with tetratricopeptide repeats 2	2.054988
23604	NM_014326	*DAPK2*	death-associated protein kinase 2	2.049421
4585	NM_018406, NM_138297, NM_004532	*MUC4*	mucin 4, cell surface associated	2.047106
4241	NM_005929, NM_033316	*MFI2*	antigen p97 (melanoma associated) identified by monoclonal antibodies 133.2 and 96.5	2.037363
10148	NM_005755	*EBI3*	Epstein-Barr virus induced 3	2.019365

80008	NM_001303228, NM_024943	*TMEM156*	transmembrane protein 156	2.019316
8638	NM_198213, NM_003733, NM_001261825	*OASL*	2'-5'-oligoadenylate synthetase-like	2.010666
7980	NM_001271003, NM_001271004, NM_006528	*TFPI2*	tissue factor pathway inhibitor 2	2.002586
		Downregulated		
6144	NM_000982	*RPL21*	ribosomal protein L21	-5.003664
6876	NM_003186, NM_001001522	*TAGLN*	transgelin	-3.989878
362	NM_001651	*AQP5*	aquaporin 5	-3.734460
3868	NM_005557	*KRT16*	keratin 16	-3.668240
101927318	NR_110589, NR_110590, NR_110591	*LOC101927318*	Not applicable	-3.316887
103344718	NM_001293171	*HOTS*	H19 Opposite Tumor Suppressor	-2.851831
1580	NM_000779, NM_001099772	*CYP4B1*	cytochrome P450, family 4, subfamily B, polypeptide 1	-2.848015
8581	NM_003695	*LY6D*	lymphocyte antigen 6 complex, locus D	-2.782002
27063	NM_014391	*ANKRD1*	ankyrin repeat domain 1 (cardiac muscle)	-2.768530
728066	NR_034169	*FAM133DP*	family with sequence similarity 133, member D, pseudogene	-2.714933
3861	NM_000526	*KRT14*	keratin 14	-2.652314
3020	NM_002107	*H3F3A*	H3 histone, family 3A	-2.519758
283120	NR_002196	*H19*	H19, imprinted maternally expressed transcript (non-protein coding)	-2.446656
4162	NM_006500	*MCAM*	melanoma cell adhesion molecule	-2.432861
100134938	NM_001114403	*UPK3BL*	uroplakin 3B-like	-2.415997
875	NM_000071, NM_001178009, NM_001178008	*CBS*	cystathionine-beta-synthase	-2.406382
29968	NM_021154, NM_058179	*PSAT1*	phosphoserine aminotransferase 1	-2.290618
479	NM_001185085, NM_001676	*ATP12A*	ATPase, H+/K+ transporting, nongastric, alpha polypeptide	-2.257360

3860	NM_002274, NM_153490	*KRT13*	keratin 13	-2.256712
440	NM_133436, NM_183356, NM_001178076, NM_001178077, NM_001178075, NM_001673	*ASNS*	asparagine synthetase (glutamine-hydrolyzing)	-2.178540
6820	NM_177973, NM_004605	*SULT2B1*	sulfotransferase family, cytosolic, 2B, member 1	-2.168962
5625	NM_001195226, NM_016335	*PRODH*	proline dehydrogenase (oxidase) 1	-2.149106
654790	NM_001102566	*PCP4L1*	Purkinje cell protein 4 like 1	-2.119195
3872	NM_000422	*KRT17*	keratin 17	-2.116836
1906	NM_001168319, NM_001955	*EDN1*	endothelin 1	-2.088337
3164	NM_002135, NM_001202233, NM_001202234, NM_173157	*NR4A1*	nuclear receptor subfamily 4 group A member 1 isoform 3	-2.049572
27122	NM_013253, NM_015881, NM_001018057	*DKK3*	dickkopf WNT signaling pathway inhibitor 3	-2.041461
3122	NM_019111	*HLA-DRA*	major histocompatibility complex, class II, DR alpha	-2.018474

[a]RefSeq mRNAs: NM_001288959, NM_005532, NM_001288958, NM_001288957, NM_001288956, NM_001288960, NM_001288954, NM_001288952, NM_001288995.
[b]RefSeq mRNAs: NM_001270927, NM_001270929, NM_001548, NM_001270928.